Joint Committee of Public Accounts and Audit

ANSWERS TO QUESTIONS ON NOTICE
Inquiry into the use and governance of artificial intelligence systems by public sector entities
Digital Transformation Agency
15 November 2024

Department/Agency: Digital Transformation Agency
Topic: Key threats to and from AI
Asked by: Senator Linda Reynolds
Type of question: Hansard
Date set by the committee for the return of answer: 3 December 2024
Question on Notice Number: DTA-UGAI-001

Number of pages: 1

Question:
Home Affairs (ASD) outlined 4 key threats to and from AI (threats from AI, threats to AI, threats via AI and accidental threats from AI) - and I'm wondering if you could take it on notice in terms of how you are dealing with those threats in terms of policies and contract management and the other frameworks you are putting in place?

Answer:

The Digital Transformation Agency (DTA) is committed to ensuring agencies are aware of their responsibilities to protect AI systems from a range of cyber threats. The DTA works closely with the Department of Home Affairs (Home Affairs) and the Australian Signals Directorate (ASD) to develop AI-related policy, assurance and technical standards that leverage and complement existing cyber security related frameworks, guidance and legislation within government.

The draft guidance for the draft Australian Government AI Assurance Framework directs agencies to implement the Protective Security Policy Framework (PSPF) and consider the recommended mitigations contained within the ASD's Engaging with Artificial Intelligence guidance.

The DTA will continue to consult with Home Affairs, ASD and other agencies to ensure alignment in the development of AI technical standards, and updates to the draft Australian Government AI Assurance Framework and the Policy for the responsible use of AI in government.

Joint Committee of Public Accounts and Audit

ANSWERS TO QUESTIONS ON NOTICE
Inquiry into the use and governance of artificial intelligence systems by public sector entities
Digital Transformation Agency
15 November 2024

Department/Agency: Digital Transformation Agency
Topic: Validating algorithms and coding for data sets
Asked by: Senator Linda Reynolds
Type of question: Hansard
Date set by the committee for the return of answer: 3 December 2024
Question on Notice Number: DTA-UGAI-002

Number of pages: 2

Question:

Can you please unpack a bit further for us is in terms of how you validate algorithms and the coding that goes into the data sets because obviously there is a significant amount of work now internationally and concern raised about where it's sourced, how it's sourced, what are the biases and other things that can go into that, and how you manage that?

Answer:

AI Assurance Framework

The Digital Transformation Agency (DTA) is conducting a pilot of the draft Australian Government AI assurance framework and supporting guidance. These materials step agencies through an AI use case impact assessment process and cover the following topic areas:

- **Fairness:** ensuring agencies have a clear definition of what constitutes a fair outcome in the context of their AI use case and processes to measure fairness, to help identify and mitigate fairness risks.
- **Data suitability:** considering the suitability of the data used to operate, train and validate an AI system, including its provenance, lineage and volume, and whether it includes any Indigenous data, in line with the Framework for Governance of Indigenous Data.
- **Testing:** developing a testing plan to help validate an AI system is performing as intended. This plan should include specific, objective and verifiable acceptance criteria, as well as testing against test datasets to look for biases or possible unintended outcomes before real-world deployment.
- **Monitoring:** establishing a plan to actively monitor and evaluate AI system performance, including to identify errors, biases and anomalies that can impact safety and reliability, such as concept or data drift.

- **Preparedness to intervene or disengage:** ensuring clear processes are in place so that anyone interacting with an AI system can report and escalate issues to a person with the authority to quickly intervene in and/or safely disengage the AI system.
- **Privacy protection and security:** minimising and protecting personal information, assessing privacy impacts and ensuring alignment with the Australian Privacy Principles and Protective Security Policy Framework.
- **Human-centred values:** ensuring the AI system aligns with human rights obligations, and incorporating diverse perspectives at all stages of the AI system lifecycle, to promote inclusion and mitigate biases.

The draft framework provides a structure for agencies to document risks and the actions they are taking to manage them. The supporting guidance provides further advice on completing these assessments, and points to trusted external resources (such as the National AI Centre's Implementing Australia's AI Ethics Principles).

Findings from the pilot will inform advice to government on the next steps for AI assurance.

Technical Standards

The DTA is developing technical standards and advice that covers an AI lifecycle (based on the National AI Centre's model), providing guidance to agencies on how to build safe and responsible use into every stage of the technology. This approach will help identify and mitigate the risk of unintended bias within existing and developing AI models as well as within the associated data needed to build, train and validate or ground AI models. This will include technical guidance on data selection, model selection and training, and evaluation techniques to ensure safe and responsible AI solutions in the Australian Public Service.

Joint Committee of Public Accounts and Audit

ANSWERS TO QUESTIONS ON NOTICE
Inquiry into the use and governance of artificial intelligence systems by public sector entities
Digital Transformation Agency
15 November 2024

Department/Agency: Digital Transformation Agency
Topic: Visibility and Transparency in the use of AI
Asked by: Senator Linda Reynolds
Type of question: Hansard
Date set by the committee for the return of answer: 3 December 2024
Question on Notice Number: DTA-UGAI-003

Number of pages: 2

Question:

How do we best develop this framework and providing visibility and transparency as I said at the beginning to all Australians because I think we do have a duty to Australians so that they understand not only how their data is being used, but how their decisions are being made, and what information or what letters or anything that they get may be AI informed or generated?

Answer:

The Digital Transformation Agency (DTA) has released the Policy for the responsible use of AI in government, which came into effect 1 September 2024. Compliance with the Policy is mandatory for all Non-corporate Commonwealth entities, except for the Defence portfolio and national intelligence community agencies.

The Policy sets the requirement that agencies will need to publish an AI transparency statement by 28 February 2025. Agencies must apply the Standard for AI transparency statements when creating their statements to ensure they provide a minimum level of information about their AI use, including:

- why the agency uses or is considering using AI
- AI uses by classification
- measures to monitor the effectiveness of deployed AI systems
- compliance with applicable legislation and regulation
- efforts to identify and protect the public against negative impact

The classification system for AI uses represents how AI is commonly used in government and the domains where they are applied. Agencies will list both the usage patterns and domains which apply to their use of AI in their transparency statements.

The DTA is also piloting the draft Australian Government AI Assurance Framework. The framework outlines additional transparency mechanisms for agencies to consider in the context of assessing their AI use case against Australia's AI Ethics principles. The outcome of this pilot will inform next steps for the Policy for the responsible use of AI in government.